DOCUMENT RESUME

ED 389 743 TM 024 376

AUTHOR Kaiser, Paul D.; Brull, Harry
TITLE New Stuff in I/O (In-Baskets and Orals). The Development, Administration and Scoring of In-Baskets and Orals for the New York State Correction Captain Examination.
PUB DATE 29 Jun 94
NOTE 21p.; Paper presented at the International Personnel Management Association Assessment Council Conference (Charleston, SC, June 29, 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Competitive Selection; Correctional Rehabilitation; Decision Making; Educational Innovation; Evaluators; Interrater Reliability; *Personnel Selection; *Scoring; Simulation; *Test Construction; Testing; Test Items; Test Results
IDENTIFIERS *In Basket Simulation; *Oral Examinations; Performance Based Evaluation

ABSTRACT

The design, administration, scoring, and results of the 1993 New York State Correctional Captain Examination are described. The examination was administered to 405 candidates. As in previous Sergeant and Lieutenant examinations, candidates also completed latent image written simulation problems and open/closed book multiple choice test components. New approaches to constructing the examination permitted the inclusion of in-baskets and orals. The in-basket items required candidates to handle relatively complex situations. Candidates were then asked specific questions about their decisions. Scoring was by rater pairs formed of a civil service examiner and a corrections subject matter expert. Each rater pair scored a small subset of items for each candidate, becoming item experts who could evaluate a small set of responses rapidly and consistently. The inclusion of oral interviews provided other challenges. The interviews were conducted by a trained civil service examiner who asked standardized questions. Assessors, Department of Corrections personnel above the rank of captain, evaluated responses on a 9-point scale. Both in-basket and oral processes proved to be efficient and economical. Two appendixes contain an examination type synopsis and the history of litigation about corrections department examinations. (Contains 10 tables.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 389 743

NEW STUFF IN I/O (IN-BASKETS AND ORALS)

The Development, Administration and Scoring of In-Baskets and Orals for the New York State Correction Captain Examination

Presented at:

The
International Personnel Management
Association Assessment Council Conference
Charleston, South Carolina
June 29, 1994

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
HARRY BRULL
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

Paul D. Kaiser
Principal Examiner
New York State
Department of Civil Service
Albany, New York

Harry Brull
Vice President
Public Sector Services
Personnel Decisions, Inc.
Minneapolis, Minnesota

BEST COPY AVAILABLE

TM024376

Overview of the Presentation

The challenge was straightforward. How can In-baskets and Orals be economically incorporated into an exam plan with more than 400 candidates? What new approaches were necessary to improve the efficiency and effectiveness of two potentially powerful, but traditionally expensive and sometimes unreliable, examination procedures?

This presentation will describe the design, administration, scoring, and results of the 1993 New York State Correctional Captain examination. The Captain examination was administered to 405 applicants. As in the previous Sergeant and Lieutenant examinations, candidates also completed latent image written simulation problems and open/closed book multiple choice components. Some new approaches permitted the inclusion of In-baskets and Orals, even with an applicant pool of this size.

New York State's Correctional Services Testing Program

Background

There are four primary titles in the security promotion field in DOCS. They are: Correction Officer, Sergeant, Lieutenant and Captain. From 1972 until 1987, every promotional examination given for these titles had been subject to Federal and State Court scrutiny. Because of the litigation, the only promotional title for which an examination was held between 1972 and 1980 was Sergeant; that examination was developed with the direct involvement of the Federal Court and was pronounced valid and fair by the Court and EEOC.

Examination Innovations and History

Because of the critical nature of positions in the correction security field and the high potential for litigation, New York State has employed independent consultants to guide the development and scoring of each promotional examination given since 1984. In each instance the consultant was a nationally recognized leader in employee selection. In some cases (e.g., when mandated by existing consent decrees) we employed an additional consultant designated by the plaintiffs.

In addition to ensuring that proper job analysis studies linked the examinations to critical job duties and that the examinations' results were evaluated for adverse impact and statistical soundness, the use of these consultants resulted in a number of innovative approaches to examination content, administration and scoring. These include: study guides, oral interviews, video tests, written simulations, open book tests, memory tests, in-basket exercises, performance evaluation and criterion validity studies, sample tests, score adjustments to reflect on-the-job performance and zone scoring. A more detailed description of these practices is presented in Appendix A.

Litigation History

A succinct overview of the various legal actions surrounding examinations administered for security service positions in the Department of Correctional Services is presented in Appendix B. The examination for Correction Captain had a very high potential for Federal and State court intervention. Since the 1970s all promotion examinations for correction security positions have had court involvement - this litigation ranged from timeliness and "score adjustment" issues to Title VII suits. The previous examination was administered pursuant to a consent decree. At the time of the present examination over half the positions were refilled provisionally -- many of these provisionals were protected calls individuals.

The 1993 Correction Captain Selection Plan

The examination components included in the 1993 Captain's examination were described in a Study Guide provided to all applicants six weeks prior to the test date. This study guide presented candidates with information on how to prepare for and take the examination as well as practice tests to assist in their preparation. Briefly, the examination included four equally weighted parts:

1. **A Multiple Choice Test**
   a) Open Book knowledge application - 45 items
   b) "Memory" knowledge application - 15 items
   c) Assessment of written communications ability - 30 items

2. **Incident Simulation Exercises** - This component consisted of three latent-image simulation problems designed to assess candidates' knowledge, skills and ability to solve problems, gather information and make decisions in correctional security situations. Each of the problems centered on a critical aspect of the captain's position. Problem One: an emergency situation involving inmate movement and control-potential inmate violence and conflict. Problem two: supervision of command staff -- the Lieutenants and Sergeants reporting to the Captain. Problem three: Oversight of an investigation involving possible inmate (and staff) criminal activity.

3. **In-Basket Exercises** - This component of the examination was designed to assess the candidates' ability to deal with the day-to-day issues and problems which confront Correction Captains. Candidates were presented with material in the same way that it 'hits' the Captains desk - in the form of memos, telephone messages, notes from superiors, colleagues and subordinates, etc. Candidates were required to respond to each item presented in free-response written form.

4. **Structured Oral Interview** - This exercise was designed to assess candidates' performance in six dimensions: Leadership, Interpersonal Skills, Conscientiousness and Dedication, Fairness and Self-Awareness, Handling Stress and Oral Communication Skills.

THE IN-BASKET EXERCISE

A review of the literature and experience from present practice suggests that traditional In-baskets suffer from a variety of problems, particularly in scoring. The typical concept of multi-dimensional (decision-making, leadership, administrative skills, etc.) ratings has yielded results of questionable reliability and has required large amount of scoring effort. The size of the candidate pool, the scarcity of resources and the litigious environment in New York necessitated fresh strategies.

The new approaches to the In-basket began with changes in item design. Instead of the usual inclusion of many "fluff" items, the fifteen-item In-basket required candidates to handle relatively complex situations -- all of which required significant action and decisions. Candidates were then directed in their response to answer specific questions regarding each item, such as "what exactly would you now do?" and "whom would you inform?", etc.

The scoring design and rating process represented the greatest departure from typical practice. Rater pairs were formed consisting of a civil service examiner and a subject matter expert (Correction Captain or Deputy Superintendent). Rating scales were constructed for each item and further refined by the group of raters. Each rater pair then scored a small subset of items (2 or 3) for all candidates. Each pair thus became "item experts" who could rapidly compare a candidate's response against a very sharply defined criteria of poor, acceptable, and better-than-average response. The combination of item design, candidate response format, and revised scoring process dramatically reduced the time required for scoring.

The In-basket item thus became the unit of measurement. A competence score (ranging from 0 to 5) was assigned to each item by comparing the candidate's response against the rating scale. Efficiency, or the traditional In-basket administrative score, was computed by giving an additional point for completing an item at the acceptable or better level. Ten of the fifteen items were previously identified as "high-priority". Each received an additional point if successfully completed.

Two Sample items, the item analysis and the rating standards are presented below. Tables One through Five summarize the psychometric properties of this modified In-basket.

Sample in-Basket Items

ITEM 8 - PHONE CALL ABOUT BLACK EYE

TELEPHONE CALL

The Watch Commander calls and says:

> "I have the Draft Processing Sergeant Sproule on the other line. The transportation bus just dropped off five inmates and one has a serious black eye but there is no documentation on how he got it. When Sproule asked him how he got the black eye, the inmate said that he 'got beaten up at the other facility.' The bus has already left so the Sergeant didn't get to talk with the transporting officers. The Sergeant is asking what he should do. What do you want me to tell him?"

Assume the Watch Commander is still on the line.

Question:

What would you tell the Watch Commander to do?

* * * * *

SCORE REPORT ITEM 8
PRIORITY RATING - HIGH

Score	Frequency	Percent
1.00	36	8.8
2.00	123	30.1
3.00	172	42.1
4.00	55	13.4
5.00	1	.2
.00	18	5.4

IN-BASKET ITEM

ITEM 10 - NOTE CONCERNING RELIGIOUS FUNCTION

At 16:45 p.m. your secretary passes you the following note.

The Watch Commander just called. Apparently, the Program Sergeant just found out that 70 Muslim inmates are scheduled for a religious function in the chapel at 17:30 this evening. There is no scheduled coverage for that area. Also, the minister who is to lead the program just called in sick and will not be available. The Watch Commander wants to know if we should cancel the program. He says it would sure make things easier.

Question:

What actions would you take in response to this situation?

* * * * *

SCORE REPORT ITEM 10

PRIORITY RATING - HIGH

Score	Frequency	Percent
1.00	49	12.0
2.00	72	17.6
3.00	157	38.4
4.00	46	11.2
5.00	1	.2
.00	80	20.0

RATING STANDARDS - Item 8-Phone Call about Black Eye

Superior (5) - To receive an **superior rating of 5**, a candidate is required to include references to **EACH OF** the following eight actions.

- Direct that the inmate be taken to medical for evaluation.
- Direct that the inmate's injury be photographed.
- Have the injured inmate interviewed to get details.
- Have the other four inmates interviewed to see if they know anything.
- Have the facility that the inmate came from called to see if they know anything about the inmate's injury
- Attempt to determine how the inmate was processed into the facility (off the bus) without appropriate medical documentation check concerning the injury.
- Direct that an entry be made in an appropriate log documenting the situation with as much detail as possible.
- If investigation warrants, direct that an U.I. be prepared and CCC or Inspector General's Office be informed

Highly Satisfactory (4) - To receive a **highly satisfactory rating of 4**, a candidate must include references to **EACH OF** the following <u>four</u> actions.

- Direct that the inmate be taken to medical for evaluation.
- Direct that the inmate's injury be photographed.
- Have the injured inmate interviewed to get details.
- Have the other four inmates interviewed to see if they know anything.

And Any One of the Following Actions

- Have the facility that the inmate came from called to see if they know anything about the inmate's injury
- Attempt to determine how the inmate was processed into the facility (off the bus) without appropriate medical documentation check concerning the injury.
- Direct that an entry be made in an appropriate log documenting the situation with as much detail as possible.
- If investigation warrants, direct that an U.I. be prepared and CCC or Inspector General's Office be informed

Satisfactory (3) - To receive a **satisfactory rating of 3**, a candidate must include references to **EACH OF** the following three actions.

- Direct that the inmate be taken to medical for evaluation.
- Direct that the inmate's injury be photographed.
- Have the injured inmate interviewed to get details.

And Either

- Have the other four inmates interviewed to see if they know anything.

- or -

- Have the facility that the inmate came from called to see if they know anything about the inmate's injury

Below satisfactory (2) - To receive a **below satisfactory rating of (2)**, a candidate's response must include references to **BOTH** of the actions listed below:

- Direct that the inmate be taken to medical for evaluation.
- Direct that the inmate's injury be photographed.

Unsatisfactory (1) - To receive an **unsatisfactory rating of (1)**, a candidate's response must include references to **at least one** of the actions listed below:

- Direct that the inmate be taken to medical for evaluation.
- Direct that the inmate's injury be photographed.
- Have the injured inmate interviewed to get details.
- Have the other four inmates interviewed to see if they know anything.
- Have the facility that the inmate came from called to see if they know anything about the inmate's injury
- Attempt to determine how the inmate was processed into the facility (off the bus) without appropriate medical documentation check concerning the injury.
- Direct that an entry be made in an appropriate log documenting the situation with as much detail as possible.
- If investigation warrants, direct that an U.I. be prepared and CCC or Inspector General's Office be informed

Totally unsatisfactory (0) - To receive a **totally unsatisfactory rating of (0)** a candidate's response is either missing, unreadable or not referencing any of the actions listed above.

INAPPROPRIATE ACTIONS

Any one of the following steps would be inappropriate under the circumstances presented; therefore if included in a candidate response will result in a quantitative lowering of a candidate's score:

- Delay medical evaluation.
- Accept and act upon inmate's complaint without investigation.

RATING STANDARDS - **Item 10 Religious Function**

Superior (5) - To receive a **superior rating of 5**, a candidate is required to include references to **EACH OF** the following **three** steps.

- Attempt to make arrangements to hold the function by locating another who could lead the service.
- Find other means of providing necessary security coverage.
- Attempt to determine why the program had not been properly scheduled with security staff.

and one of the following actions

- Counsel the Watch Commander for the "suggestion" that the function should be canceled to "make things easier..."
- Consider the effects of other simultaneous or recent events -- such as the Mess Hall Fight or the Medical Trip.

Note: To receive a Superior rating, a candidate must make every attempt to hold the function. However, if the candidate indicated that the above actions were not successful and that the function must be postponed, then the candidate must also include the statement that the captain, not someone else, would inform the inmates and that they would be informed of the attempts made to move forward with the important function.

Highly Satisfactory (4) - To receive a **highly satisfactory rating of 4**, a candidate must include references to **EACH OF** the following two actions.

- Attempt to make arrangements to hold the function by locating another who could lead the service.

- Find other means of providing necessary security coverage.

and one of the following actions

- Attempt to determine why the program had not been properly scheduled with security staff.
- Counsel the Watch Commander for the "suggestion" that the function should be canceled to "make things easier..."
- Consider the effects of other simultaneous or recent events -- such as the Mess Hall Fight or the Medical Trip.

Satisfactory (3) - To receive a **satisfactory rating of 3**, a candidate must include references to **both of** the following two actions which would serve to facilitate a positive outcome toward holding the function.

- Attempt to make arrangements to hold the function by locating another who could lead the service.
- Find other means of providing necessary security coverage.

Note: The only time that an answer was rated satisfactory WITHOUT containing **BOTH** of the above points was when security coverage WAS found and recommendation was made to HOLD the service. In this situation, it is also assumed that the inmates would hold the service even without a religious leader--which is permissible.

Below satisfactory (2) - To receive an **below satisfactory rating of (2)** a candidate's response must include references to **both** of the actions listed below:

- Attempt to make arrangements to hold the function by locating another who could lead the service.
- Find other means of providing necessary security coverage.

And

- Recommend that the function be cancelled even through both the above actions were taken.

Unsatisfactory (1) - To receive an **unsatisfactory rating of (1)** a candidate's response must include references to **any of the** actions listed below:

- Attempt to determine why the program had not been properly scheduled with security staff.
- Counsel the Watch Commander for the "suggestion" that the function should be canceled to "make things easier..."
- Consider the effects of other simultaneous or recent events -- such as the Mess Hall Fight or the Medical Trip.

Totally unsatisfactory (0) - To receive a **total'y unsatisfactory rating of (0)** a candidate's response is either missing, unreadable or not referencing any of the steps listed above.

APPROPRIATE ACTIONS

Any one of the following steps would be inappropriate under the circumstances presented; therefore if included in a candidate response will result in a quantitative lowering of a candidate's score:

- Cancel the service without even having tried to find either a religious leader or security coverage.

THE ORAL INTERVIEW COMPONENT

The inclusion of Orals presented another challenge. The last use of an oral examination in the Correctional Promotion series required the use of "outside" raters who were flown in, at great expense, for weeks of serving as oral panel raters. Again, changes in the design, administration, and scoring resulted in a more efficient effective and economical process.

The first major decision was to use Orals to measure KSAs which are typically not attempted with large applicant pools. The New York Correction Captain Oral measured the following components:

- Leadership Skills
- Interpersonal Ability
- Handling Stress
- Conscientiousness and Dedication
- Fairness and Self-awareness
- Oral Communication Skills

The oral interview was designed to measure these skills using a behavioral consistency approach. Candidates were asked to describe instances where they had demonstrated these skills in the past. These Orals are based on the premise that past behavior in similar circumstances will predict future behavior. This approach made particular sense because, at least at the macro level, all candidates had spent many years in the correctional system moving through the ranks of Officer, Sergeant, and Lieutenant.

The Orals were conducted by a trained civil service examiner who asked a series of standardized opening questions and then probed the candidates' responses in order to obtain detailed descriptions of the candidates' actions.

For each of the competencies standard opening questions, suggested follow-up probes, and scoring anchors were developed. A sample of the standard opening question and possible follow-up probes are presented below.

Standard Opening Question:	Tell me about the last time you helped out another correction co-worker.
Possible Follow-up Probes:	* What was the situation? * What exactly did you do?
Standard Opening Question:	Tell me about the person with whom you have found it most difficult to work.
Possible Follow-up Probes:	* What made this person so difficult? * What approaches have you used with this person? * How successful have you been at improving the relationship?

The assessors for the Oral interview were Department of Corrections personnel at ranks above that of Correction Captain. Prior to the scoring of the examination, candidates received the list of all departmental raters, and raters received a corresponding list of candidates. Each was instructed to indicate which individuals should not be paired as evaluator/candidate. These were then honored in assignments with anonymity further ensured by identification of candidates only by code number.

Department of Corrections evaluators were trained in the same session as the Department of Civil Services interviewers. Candidate responses were evaluated on a 9-point scale against established criteria. an example of these criteria are presented below.

RATING CRITERIA EXAMPLES

Examples of low-rated responses:

* Can't think of any examples of helping others or describes trivial situation
* Creates "win/lose" situation
* Approach is likely to turn others off
* Appeals to higher authority to enlist cooperation
* Blames others for problems in coordination -- takes no or ineffective action to facilitate cooperation or get others' "buy-in"
* Provides example from many years ago

Examples of high-rated responses:

* Approach fosters teamwork
* Demonstrates ability and eagerness to develop positive relations
* Makes large sacrifice to help out an other
* Creates "win/win" situation
* Takes responsibility for problems in coordination, and takes effective action to remove obstacles and enlist "buy-in"

Following the administration of the oral interviews, the audio cassettes were distributed, by mail, to the appropriate evaluators. Independent evaluations were made and scores called in to a central number. Again, as in the in-basket, score discrepancies larger than a pre-determined distance were returned to evaluators for further discussion and consensus.

This process saved an estimated tens of thousands of dollars in rater expenses. In addition, it allowed for truly "blind" ratings of candidates' responses. A significant benefit of this procedure is that individuals who have served in the position and currently supervise Correction Captains were able to rate the candidate's performance. Thus, rather than importing experts from other jurisdictions, the individuals who know most about the job felt comfortable in rating the candidates.

The design of the Oral allows for inter-rater reliability estimates as well as comparisons of performance on the Oral with the other examination components. Tables Six A, B, and C describe the mean scores for the interview by ethnicity, the mean scores by component and the intercorrelation between components. Reliabilities reported in table Six D represent the independent judgments for each of the two-person panels who evaluated approximately 40 candidates each. Ratings from the Civil Service interviewers, although not used in the scoring process, were also collected.

SUMMARY

Both the In-basket and the Oral Interview processes proved to be efficient and economical. Once the In-basket scoring standards were finalized, the actual rating process for the 405 candidates required approximately 250 rater-hours to complete. On average, each In-basket was scored in about 40 minutes by **both** raters. This compares very favorably with some of the estimates cited in the literature of 2 to 5 hours of scoring time per In-basket. The average inter-rater correlation of .82 is within the standards set by professional practice. The In-basket had less adverse impact against Blacks than either the multiple choice or simulation component. Candidate acceptance of the In-basket was high; none of the scoring standards were challenged by the candidates during the appeal sessions, and all of the ratings assigned to candidates remained unchanged after appeal.

The Oral Interview process was implemented at minimal expense - the only direct costs were the one-day rater training session and the cassette tapes. Overall savings, when contrasted with the more traditional three-person Structured Oral Broad, are estimated to be in the $90,000+ range. The assessment literature suggests that the behavioral consistency model employed has the greatest potential for the highest validity. The use of blind rating panels allowed the ratings to be done by those subject matter experts who know most about the job. The inter-rater correlations ranging from .53 to .89 sufficiently demonstrate the reliability of the process.

TABLE ONE

FREQUENCY DISTRIBUTION OF FINAL IN-BASKET SCORES

Score	Frequency	Cum Percent
65+	3	.7
60-64	4	1.7
55-59	13	4.9
50-54	27	11.6
45-49	47	23.2
40-44	59	37.8
35-39	57	51.9
30-34	81	71.9
25-29	48	83.7
20-24	37	92.8
0-19	29	100.0
TOTAL	405	100.0

Mean = 35.8
Std Dev = 11.0

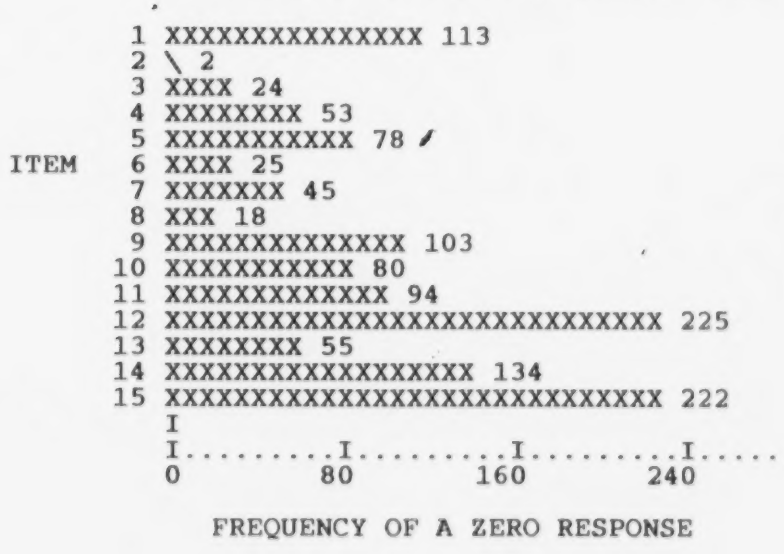
TABLE TWO A

DISTRIBUTION OF NO RESPONSE BY ITEM

Item	Frequency of Zero	Percent of Candidates
1	113	28.0
2	2	0.0
3	24	6.0
4	53	13.0
5	78	19.0
6	25	6.0
7	45	11.0
8	18	4.0
9	103	25.0
10	80	20.0
11	94	3.0
12	225	56.0
13	55	14.0
14	134	33.0
15	222	55.0

TABLE TWO B

HISTOGRAM OF NO RESPONSE BY ITEM



FREQUENCY OF A ZERO RESPONSE

TABLE THREE

IN-BASKET ITEM MEANS - RAW/RAW WITH NON-RESONANTS DELETED/ RAW WITH PRIORITY & EFFICIENCY POINTS

ITEM	MEAN W/ ZEROS	MEAN W/O ZEROS	MEAN OF FINAL SCORE
1	1.28	1.78	1.41
2	2.35	2.36	3.04
3	2.26	2.40	3.01
4	1.65	1.90	2.06
5	1.68	2.09	1.96
6	2.35	2.51	3.28
7	2.52	2.83	3.10
8	2.53	2.64	3.65
9	1.78	2.39	2.40
10	2.11	2.62	3.11
11	2.13	2.77	2.59
12	.94	2.11	1.21
13	1.66	1.92	1.98
14	1.21	1.80	1.32
15	1.17	2.59	1.63
TOTAL	27.62	34.71	35.75
Coefficient Alpha for Final Score = .46			

TABLE FOUR

IN-BASKET FINAL SCORES BY ETHNICITY

	Mean	Std Dev	Cases
WHITE	37.1	11.0	340
BLACK	27.4	6.4	51
HISPANIC	34.0	11.4	12
AM IND	39.0	.0	1
TOTAL	35.7	11.8	405

TABLE FIVE

IN-BASKET W/O ZEROS MEANS AND INTERRATER CORRELATIONS

Item	Means		Γ_{AB}
	Rater A	Rater B	
1	1.89	1.72	.80
2	2.35	2.35	.87
3	2.26	2.33	.67
4	1.99	1.90	.76
5	2.02	2.15	.69
6	2.50	2.51	.93
7	2.85	2.78	.80
8	2.59	2.62	.72
9	2.44	2.39	.81
10	2.59	2.64	.90
11	2.72	2.71	.77
12	2.09	2.12	.83
13	1.92	2.03	.66
14	1.81	1.86	.88
15	2.51	2.58	.90
Total	34.53	34.69	.82*

*Average inter-correlation

TABLE SIX A

ORAL TEST COMPONENT RESULTS
ORAL CONSENSUS TOTAL SCORES BY RACE

	Mean	Std Dev	Cases
WHITE	33.6	6.3	338
BLACK	32.1	6.8	50
HISPANIC	29.0	7.9	12
AM IND	33.0	.0	1
TOTALS	33.3	6.4	401

TABLE SIX B

MEAN SCORE BY DIMENSION

Dimension	Mean	Std Dev	Cases
LEADERSHIP	5.5	1.2	401
INTERPERSONAL	5.5	1.1	401
CONSCIENTIOUS	5.4	1.2	401
FAIRNESS	5.4	1.2	401
STRESS	5.4	1.2	401
ORAL SKILL	6.0	1.4	401

TABLE SIX C

CORRELATION MATRIX BY DIMENSION

	LEAD	INTER	CONSC	FAIR	STRES	ORAL
LEAD	1.0	.69	.62	.61	.55	.59
INTER		1.00	.66	.70	.63	.67
CONSC			1.00	.73	.68	.64
FAIR				1.00	.74	.72
STRES					1.00	.77
ORAL						1.00

* All correlations are significant

TABLE SIX D

ORAL RATER MEANS AND INTER-RATER CORRELATIONS

	Means		
Board	Rater A	Rater B	Γ_{AB}
1	34.90	35.81	.75
2	28.61	31.70	.71
3	30.97	30.73	.83
4	34.69	33.01	.53
5	35.54	35.43	.86
6	29.98	29.08	.78
7	36.28	37.08	.78
8	29.00	30.20	.89

TABLE SEVEN

MULTIPLE CHOICE COMPONENT RESULTS

MEMORY QUESTIONS (15 Questions weighted 1.0)

	Mean	Std Dev	Cases
WHITE	10.4	1.9	340
BLACK	9.4	2.3	52
HISPANIC	9.1	2.2	12
AM IND	12.0	.0	1
TOTALS	10.3	2.0	405

VERBAL QUESTIONS - Reading Comp & Report Prep. (30 Questions weighted .5)

	Mean	Std Dev	Cases
WHITE	8.1	2.3	340
BLACK	6.4	1.9	52
HISPANIC	6.2	1.9	12
AM IND	11.5	.0	1
TOTALS	7.8	2.3	405

OPENBOOK QUESTIONS - Use References to answer (45 Questions weighted 1.0)

	Mean	Std Dev	Cases
WHITE	33.8	5.6	340
BLACK	28.4	6.5	52
HISPANIC	28.7	7.0	12
AM IND	27.0	.0	1
TOTALS	32.9	6.1	405

TOTAL MULTIPLE CHOICE TEST

	Mean	Std Dev	Cases
WHITE	52.3	7.7	340
BLACK	44.2	8.3	52
HISPANIC	44.0	9.3	12
AM IND	50.5	.0	1
TOTALS	51.0	8.3	405

Total Multiple-Choice Test Coefficient Alpha = .77

TABLE EIGHT A

SIMULATION TEST COMPONENT RESULTS

PROBLEM 1

	Mean	Std Dev	Cases
WHITE	18.4	4.7	337
BLACK	16.2	6.0	51
HISPANIC	19.3	4.5	12
AM IND	15.0	.0	1
TOTALS	18.1	4.9	401

PROBLEM 2

	Mean	Std Dev	Cases
WHITE	22.8	4.4	337
BLACK	19.6	6.0	51
HISPANIC	19.8	4.2	12
AM IND	20.0	.0	1
TOTALS	22.3	4.8	401

PROBLEM 3

	Mean	Std Dev	Cases
WHITE	30.8	5.5	337
BLACK	25.6	7.3	51
HISPANIC	26.3	7.1	12
AM IND	40.0	.0	1
TOTALS	30.0	6.1	401

SIMULATION TOTAL TEST SCORE

	Mean	Std Dev	Cases
WHITE	71.9	9.9	337
BLACK	61.2	14.7	51
HISPANIC	65.3	11.5	12
AM IND	75.0	.0	1
TOTALS	70.4	11.3	401

TABLE EIGHT B

SIMULATION PROBLEM CORRELATION MATRIX

Problem	SIM1	SIM2	SIM3
SIM1	1.0000	.1800**	.1985**
SIM2	.1800**	1.0000	.3813**
SIM3	.1985**	.3813**	1.0000

TABLE NINE

TOTAL TEST SUMMARY STATISTICS
ADVERSE IMPACT ANALYSIS

MULTIPLE CHOICE COMPONENT

	Mean	Std Dev	Cases	Diff
WHITE	.16	.92	340	-
BLACK	-.82	1.00	52	.98
HISPANIC	-.84	1.11	12	1.00
MALE	.00	1.00	391	-
FEMALE	-.12	.90	14	.12
TOTALS	0.00	1.00	405	

SIMULATION COMPONENT

	Mean	Std Dev	Cases	Diff
WHITE	.13	.88	340	-
BLACK	-.77	1.30	52	.90
HISPANIC	-.43	1.00	12	.56
MALE	.02	1.00	391	-
FEMALE	-.46	.88	14	.48
TOTALS	.00	1.00	405	

ORAL COMPONENT

	Mean	Std Dev	Cases	Diff
WHITE	.05	.98	338	-
BLACK	-.18	1.06	50	.23
HISPANIC	-.66	1.24	12	.71
MALE	.00	1.00	387	-
FEMALE	.00	1.10	14	.00
TOTALS	.00	1.00	401	

IN-BASKET COMPONENT

	Mean	Std Dev	Cases	Diff
WHITE	.12	1.01	340	-
BLACK	-.76	.58	52	.88
HISPANIC	-.16	1.04	12	.28
MALE	.02	1.00	391	-
FEMALE	-.44	.97	14	.46
TOTALS	.00	1.00	405	

TABLE TEN

CORRELATION MATRIX BY TEST TYPE

	MC	SIM	ORAL	IN-BASKET
MC	1.0000	.5023**	.0993	.3176**
SIM	.5023**	1.0000	.0590	.2696**
ORAL	.0993	.0590	1.0000	.2067**
IN-BASKET	.3176**	.2696**	.2067**	1.0000

APPENDIX A
EXAMINATION TYPE SYNOPSIS

Study Guides: Provide candidates with complete information on the content and nature of the examination before the test. Include tips on taking the test, lists of important rules and regulations, practice tests, etc.

Written Simulations: These latent-image candidate-interactive tests are designed to evaluate the candidates' higher level decision-making, problem solving and analytical skills. Unique in that they provide feedback to the candidates while they are taking the test, they place test takers in real life situations and require that they resolve problems typically encountered in correction security positions.

Open Book Tests: Since on the job an incumbent could refer to appropriate reference material to handle certain situations, the open book examination uses the same approach. At the test site candidates are provided with copies of the rules, regulations, directives, employee and inmate handbooks, etc., found on the job and may use this material to decide upon the best answer to the questions.

Memory Tests: These questions cover subject matter which candidates are required to know from memory and cannot "look up" on the job. Copies of the critical rules, regulation, directives, etc., are provided to candidates in the study guide to allow them to prepare for the test.

In-Basket: In order to more accurately evaluate the ability of candidates to deal with situations and problems of managing one's time while dealing with important issues, we have developed the use of In-Basket Written Exercises. Using this format, candidates are presented with material in much the same way that it 'hits' the desks of incumbents in the position for which the examination is being developed--memos, telephone messages, notes from superiors, colleagues and subordinates, routine (and non-routine) department and facility forms, directives and paperwork, letters, etc. The candidates are then required to evaluate the priority of the problems in consideration of limited time and resources and to prepare brief statements in response to specific questions asked concerning how they would deal with specific issues or how they would handle specific problems in relation to the materials in their in-basket.

Video Tests: This type of test uses video to present candidates with real-life situations in a non-written format. It allows for the presentation of complex scenarios which typically would require extensive written description. It tests for observational skills (including detection of non-verbal stimuli) as well as decision-making, problem solving and human interaction skills without placing undue emphasis on written comprehension.

Oral Interviews: Test for candidates' skill and ability to solve problems, reason clearly, exercise sound judgment, and communicate effectively in a correctional setting. Candidates are presented with a variety of problems to which they respond orally.

Performance Evaluation and Criterion Validity Studies: Data on candidate and incumbent performance was collected as part of the candidates' final score on the examination as a criteria for evaluating the effectiveness of the examination and a gauge for determining the fairness of the examinations' results for various ethnic groups.

Score Adjustments: Based upon criterion performance information, the scores of protected class candidates were adjusted upward so that the differences in observed test performance between minority and non-minority candidates reflected the actual job performance expectations for the different groups.

Zone Scoring: Following the Court's guidance found in the consent decree (this settlement resulted in a ratio minority appointments) that we consider alternative scoring strategies, and on the advice of the plaintiff's designated consultant in this case, we have zone scored the results of most of the examinations in DOCS from 1983 forward. Zone scoring is the grouping of different candidate raw scores within a narrow range into the same final score.

APPENDIX B
LITIGATION HISTORY

1972 October - **Correction Sergeant**: Immediately after the eligible list for this examination was established suit was brought challenging the results as adversely affecting minority candidates. Judge ruled the examination invalid due to insufficient job analysis and other defects. He ordered that an examination for Correction Sergeant--which would be criterion validated as far as possible--be developed under the supervision of the court.

1976 and 1978 October - **Interim and final court supervised Correction Sergeant**: This selection procedure took over four years to develop and included both written examination and performance evaluation components. Post-test analysis resulted in a score adjustment based upon race. The court, with the concurrence of EEOC, pronounced the examination valid and fair.

Circa 1980 - **Correction Lieutenant and Captain**: The union filed a "timeliness" suit demanding that examinations for these titles be administered. The court ordered that examinations be developed.

1981 October - **Correction Lieutenant**: The eligible list for this examination was established in December 1981. Suit was immediately brought alleging the results were discriminatory against minority candidates. The case was settled in September 1982. Provisions of the settlement included: (1) the "zone" scoring of the results of the October 1981 examination and the appointment of eligible minority candidates in a proportion of one minority to four non-minority; (2) the development, within specified time frames, of new examinations for Lieutenants and Captains; (3) the hiring of a consultant (to be designated by the plaintiffs) to advise the department in the development of new selection procedures, and in the event of adverse impact, to consider the use of "separate frequency distribution for minority and non-minority candidates."

1982 January - **Correction Captain**: Held in accordance with the dictum of the Court in response to the Union "timeliness" suit. The examination's results showed adverse impact upon minority candidates. Following the Court's suggestion in that settlement, the score results were adjusted to eliminate the adverse impact. In October 1982, the Union initiated a suit challenging the adjustment and obtained an injunction prohibiting the use of the eligible list for making permanent appointments.

1983 December - **Correction Captain**: This examination was held at the direction of the courts as mandated by settlement. The plaintiffs sought and obtained an injunction baring the use of the examination results.

1984 December - **Correction Sergeant**: This was the first non-court supervised examination for this position since 1978. Post-test analyses of candidate examination score data and minority and non-minority on-the-job performance data indicated that ethnic differences in test performance exceeded differences in job performance. An adjustment of the minority candidates' examination scores was made such that differences in test results reflected the differences in expected job performance.

1987 September - **Correction Lieutenant**: This examination was held pursuant to the consent decree. The plaintiff's designated consultant recommended the use of a "reliability based band-width" scoring procedure. Since the Department felt this approach violated New York State law, we petitioned the Federal court to allow use of the procedure. The Court did not grant our request; the examination's results were zone scored. No subsequent litigation.

1989 May - **Correction Sergeant**: This examination has not been challenged in the courts.